1-2014

Nonparametric Testing for Anomaly Effects in Empirical Asset Pricing Models

Sainan JIN
Singapore Management University, snjin@smu.edu.sg

Liangjun SU
Singapore Management University, ljsu@smu.edu.sg

Yonghui ZHANG
Renmin University of China

Nonparametric Testing for Anomaly Effects in Empirical Asset Pricing Models*

Sainan Jin[a], Liangjun Su[a], Yonghui Zhang[b]

[a] *School of Economics, Singapore Management University*

[b] *School of Economics, Renmin University of China*

January 18, 2014

Abstract

In this paper we propose a class of nonparametric tests for anomaly effects in empirical asset pricing models in the framework of nonparametric panel data models with interactive fixed effects. Our approach has two prominent features: one is the adoption of nonparametric component to capture the anomaly effects of some asset-specific characteristics, and the other is the flexible treatment of both observed/constructed and unobserved common factors. By estimating the unknown factors, betas, and nonparametric function simultaneously, our setup is robust to misspecification of functional form and common factors and avoids the well-known "error-in-variable" (EIV) problem associated with the commonly used two-pass (TP) procedure. We apply our method to a publicly available data set and divide the full sample into three subsamples. Our empirical results show that size and book-to-market ratio affect the excess returns of portfolios significantly for the full sample and two of the three subsamples in all five factor pricing models under investigation. In particular, the nonparametric component is significantly different from zero, meaning that the constructed common factors (e.g., small minus big (SMB) and high minus low (HML)) cannot capture all the size and book-to-market ratio effects. We also find strong evidence of nonlinearity of the anomaly effects in the Fama-French 3-factor model and the augmented 4-factor and 5-factor models in the full sample and two of the three subsamples.

Key Words: Anomaly effects; Asset pricing; CAPM; Common factors; EIV; Fama-French three-factor; Interactive fixed effects; Nonparametric panel data model; Sieve method; Specification test

*We sincerely thank Subal Kumbhakar and two anonymous referees for their many insightful comments and suggestions that lead to a substantial improvement of the presentation. The first and second authors gratefully acknowledge the Singapore Ministry of Education for Academic Research Fund under grant number MOE2012-T2-2-021. Address Correspondence to: Liangjun Su, School of Economics, Singapore Management University, 90 Stamford Road, Singapore 178903; E-mail: ljsu@smu.edu.sg, Phone: +65 6828 0386.

JEL Classifications: C14, C33, C58

1 Introduction

One of the central questions in finance is why different assets have different rates of return. The classical asset pricing theory states that the cross-section of equity returns can be explained by systematic risks (betas) or factor loadings on a set of common factors. Based on Markowitz's (1952) mean-variance analysis, Sharpe (1964) and Lintner (1965) propose the *capital asset pricing model* (CAPM). They suggest the systematic (market) risk as the only factor that affects the expected returns.

However, many empirical works indicate that the CAPM beta can not completely explain the cross-section of expected asset returns; see, e.g., Friend and Blume (1970) and Stambaugh (1982). These empirically documented cross-sectional patterns in average returns, which cannot be explained by theories, are called anomaly effects. In empirical finance, there has been a large amount of literature on how firm- or asset-specific characteristics such as leverage, past returns, dividend-yield, earnings-to-price ratio and book-to-market ratio as well as size help explain the cross-section returns; see Jegadeesh and Titman (1993) for price momentum, Chordia and Shivakumar (2002) for earning momentum, Ang, Hodrick, Xing, and Zhang (2006) for idiosyncratic volatility, and Fama and French (1992, 1996), Berk (1995), and Grauer (1999) for size and book-to-market ratio effects. To explain these anomaly effects, one natural way is to include more factors to capture the behavior of expected returns and this leads to the multifactor pricing models. One key issue for multifactor pricing models is how to select the factors. Two approaches are popular in the literature. The first approach is statistical and is motivated by the arbitrage pricing theory (APT) of Ross (1976) and Huberman (1982). For example, Lehmann and Modest (1988) use factor analysis and Connor and Korajczyk (1986, 1993) use principal component analysis (PCA) to select the factors to explain the cross-section returns. The second approach is to create factors based on asset/firm characteristics or some macroeconomic variables. For example, Fama and French (1993) use size and book-to-market ratio to form factor portfolios and Chen, Roll, and Ross (1986) specify macroeconomic variables as factors. For more discussions on factor selection, see the recent surveys on empirical asset pricing by Goyal (2012), Subrahmanyam (2010), and Jagannathan, Skoulakis, and Wang (2010).

Two common beliefs are underlying the above two approaches: one is that the beta does not change over time for each asset, and the other is that the factor structure is able to capture all the cross-section returns. However, the beta may vary over time or be determined by some asset-specific characteristics. For some recent advancement on relaxing the assumption of constant beta, see Li and Yang (2011), Ang and Kristensen (2012) and Li, Su, and Xu (2013) for time-varying beta, Connor and Linton (2007) and Connor, Hagmann, and Linton (2012) for the characteristic-based factor model. In addition, there may be some anomaly effects introduced by asset-specific characteristics that cannot be captured or completely captured by a factor structure. In some cases, the asset-specific characteristics may also affect the asset returns in a direct way and should be included in the factor pricing models explicitly. For some related empirical works, see Banz (1981) for size, Rosenberg, Reid, and Lanstein (1985) for book-to-market ratio, Basu (1977) for earnings-price ratio, Brennan, Chordia, and Subrahmanyam (1998) and Zhang (2009) for size and book-to-market ratio, Chordia, Goyal, and Shanken (2012) for size, book-to-market ratio and lagged returns. For some theoretical studies, see Jagannathan and Wang (1998) for a rigorous econometric analysis of the cross-sectional regression method when time-invariant firm characteristics are employed in the factor models and Jagannathan, Skoulakis, and Wang (2003) for discussions on time-varying firm-specific characteristics.

In this paper, we are interested in testing for anomaly effects of asset-specific characteristics in factor pricing models. The commonly used method in empirical studies is the two-pass (TP) cross-sectional regression method, first proposed by Black, Jensen, and Scholes (1972) and Fama and MacBeth (1973). The idea underlying the TP method is quite simple: sort the portfolios based on the means of individual characteristics which may be associated with average returns, estimate the betas for the portfolios, and check whether the average returns can be explained by the difference in betas. This idea gives rise to a two-stage procedure in applications. In the first stage, the asset betas are estimated by time series linear regressions of the asset returns on a set of common factors. In the second stage, run the cross sectional regression of mean returns on the betas and individual asset characteristics and test for the significance of asset-specific regressors. If the average returns cannot be accounted for only by the betas, the asset-specific characteristics should be significant and included in the factor pricing model.

The primary appeal of the TP method is its simplicity. However, there are four problems associated with the TP method. The usage of estimated betas in the second stage gives rise to the well-known error-in-variable (EIV) problem. Although the TP estimators are still consistent, the variance estimator used by Fama and MacBeth (1973) is asymptotically invalid. Many works have been done to correct this problem; see Shanken (1992), Jagannathan and Wang (1998), Cochrane (2005), and Ahn, Gadarowski, and Perez (2012), among others. The second problem arises from the adoption of constructed factors such as Fama and French's three factors and some macroeconomic factors. The omission of some important factors or inclusion of some irrelevant factors may lead to the misspecification of factor pricing models and consequently inconsistent estimation and misleading inference. The third problem is the efficiency loss in the TP method. In the second stage, only average returns are used to estimate the factor risk premium. Two possible reasons for running this regression in the second stage are convenience and lack of panel data on individual characteristics. The last problem is about the underlying assumption that the firm-specific characteristics affect the returns in a linear way. Since there is no theory on how the anomaly effects can be generated, imposing such a strong functional form assumption may lead to model misspecification.

In this paper, we propose a general framework to test for anomaly effects in asset pricing models.

Our setup is more general than the traditional linear factor pricing models in the sense that it includes both observed and unobserved common factors, and incorporates a nonparametric function of individual asset-specific characteristics. We test for the anomaly effects by checking whether the estimated nonparametric function is close to zero or not based on the recent work by Su and Zhang (2013). In the case where anomaly effects are detected, it is also worthwhile to test whether the functional form is linear or not. The tools developed in these papers can also be used to test the linearity of a functional form and they complement the test developed by Su and Lu (2013) for nonparametric panel data models with additive fixed effects. We apply our method to a publicly available data set and divide the full sample into three subsamples. Our empirical results show that size and book-to-market ratio affect the excess returns of portfolios significantly for the full sample and two of the three subsamples in all five factor pricing models under investigation. We also find strong evidence of nonlinearity of the anomaly effects in the Fama-French 3-factor model and the augmented 4-factor and 5-factor models in the full sample and two of the three subsamples.

The rest of the paper is organized as follows. In Section 2, we introduce the asset pricing models and formalize the hypotheses for testing the absence of anomaly effects. In Section 3, we give the estimation method for nonparametric panel data models with interactive fixed effects or with both observed and unobserved factors. Section 4 states the test statistics and their asymptotic properties under the null. In Section 5, we apply our tests to a portfolio data set. Section 6 concludes.

2 A general empirical asset pricing model

In this paper, we assume that the excess returns are generated as follows

$$r_{it} - r_{ft} = g(X_{it}) + \Gamma_i^{0\prime} d_t + \lambda_i^{0\prime} f_t^0 + e_{it},\ i = 1, ..., N,\ t = 1, ..., T, \tag{2.1}$$

where r_{it} is the rate of return of asset i at time t, r_{ft} is the risk-free return rate at time t, $r_{it} - r_{ft}$ is the excess rate of return, $g(\cdot)$ is an unknown smooth function, X_{it} is a $d \times 1$ vector of observed characteristics of asset i at time t, d_t is a vector of *observed* or *constructed* common factors such as the market risk in CAPM and the constructed three factors in Fama-French three-factor model, Γ_i^0 is the unobserved vector of betas of d_t, f_t^0 is a vector of *unobserved* common factors, and λ_i^0 is the vector of betas of f_t^0. Here $g(\cdot)$ is used to capture the anomaly effects that cannot be explained by a factor structure. We do not specify the functional form of $g(\cdot)$ to avoid functional form misspecification. Apparently, the model includes both observed/constructed and unobserved common factors and unifies the statistical and theoretical factor selection approaches. The setup offers a great flexibility to model financial returns. When we know in advance that some constructed factors such as the market risk in the CAPM explain the cross-section returns, these factors will be included in the modes to gain efficiency.

There are three main differences between our setup and the existing literature on testing for anomaly effects of characteristics in factor pricing models. First, unlike most factor pricing models that employ either constructed factors or extracted factors based on some statistical techniques, our model incorporates both types of factors simultaneously. It is well known that despite the existence of substantial empirical evidence that some constructed factors such as Fama-French three factors have certain power to explain the cross-sectional returns, there are still some unexplained factors that play a significant role in explaining asset returns; see Kleibergen and Zhan (2013) and Levwellen, Nagel, and Shanken (2010) for the account of effects of unexplained factors. Second, unlike most existing literature that imposes a linear function form, our model allows the anomalies to be caused by time-varying asset-specific characteristics through an unknown smooth function. The inference based on the linear model can be efficient if the linear functional form is correctly specified but is generally invalid and misleading otherwise. In sharp contrast, the use of nonparametric model rules out functional form misspecification and tends to yield robust inference. Third, like Serlenga, Shin, and Snell (2004) we use a panel approach instead of the widely applied TP method. Unlike the TP method that employs the data inefficiently, the panel approach is a much more efficient method in the use of information than either time-series regressions or cross-section regressions.

The setup in (2.1) is very general and include various empirical asset pricing models as special cases; see, e.g., the following widely-used asset pricing models in the panel framework.

Example 1. (One factor model: CAPM) When the model only includes the market risk premium, (2.1) reduces to the classical CAPM

$$r_{it} - r_{ft} = \alpha^0 + \beta_i^0 \left(r_{mt} - r_{ft}\right) + e_{it},\ e_{it} \sim \text{i.i.d.}(0, \sigma^2), \tag{2.2}$$

where $r_{mt} - r_{ft}$ is the market risk premium at time t, β_i^0 is the systematic risk of asset i which is time-invariant, and e_{it} is the pricing error. In this model, $g\left(X_{it}\right) = \alpha^0$ and the set of unobserved common factors is empty.

Example 2. (Fama-French three-factor model) Noticing that both the size and book-to-market ratio have strong explanatory power on stock or portfolio returns, Fama and French (1992) include two additional factors in the CAPM model. One is SMB, which stands for "small (market capitalization) minus big", and the other is HML, which stands for "high (book-to-market ratio) minus low". They measure the historical excess returns of small capitals over big capitals and of value stocks over growth stocks. These factors are calculated with combinations of portfolios composed by ranked stocks (book-to-market ratio ranking, capital ranking) and available historical market data. Then model (2.1) becomes

$$r_{it} - r_{ft} = \alpha^0 + \beta_i^0(r_{mt} - r_{ft}) + \beta_{SMB,i}^0 SMB_t + \beta_{HML,i}^0 HML_t + e_{it},\ e_{it} \sim \text{i.i.d.}(0, \sigma^2). \tag{2.3}$$

Here $g(X_{it}) = \alpha^0$ and the set of unobserved common factors is empty. In addition to Fama-French three factors, we can include the momentum factor to capture the tendency for the stock price to form Carhart's (1997) four-factor model. Of course we can also include the liquidity factor to capture the effect of market-wide liquidity on return to form another four-factor model considered by Pástor and Stambaugh (2003).

Example 3. (Pure unobserved factor model) In empirical applications, there is always some risk of model misspecification such as omitting some important factors. A robust factor pricing model is to allow the data to be generated from a latent factor model, i.e.,

$$r_{it} - r_{ft} = \alpha^0 + \lambda_i^{0\prime} f_t^0 + e_{it}. \tag{2.4}$$

This model is motivated from the original APT model of Ross (1976) and Huberman (1982), both of whom assume a strict factor structure for the return-generating process. Chamberlain and Rothschild (1983) study the APT implication under an approximate factor model for return-generating process. Connor and Korajczyk (1986, 1988) design a scheme of extracting factors from individual stock returns. In the model (2.4) one does not use any constructed or observed risk factors but simply lets the data determine the latent factors. The number of unobserved common factors can be chosen according to some well known information criteria; see, e.g., Bai and Ng (2002). In this sense, the model is robust to factor misspecification. Here $g(X_{it}) = \alpha^0$ and the set of observed common factors is empty.

Example 4. (Linear panel data model with observed factors) Serlenga, Shin, and Snell (2004) propose the following panel data model with observed factors

$$r_{it} - r_{ft} = \theta^{0\prime} X_{it} + \Gamma_i^{0\prime} d_t + u_{it}, \tag{2.5}$$

where X_{it} is a $d \times 1$ vector of regressors, $u_{it} = \alpha_i^0 + e_{it}$, and α_i^0 and e_{it} represent the individual effects and idiosyncratic error terms, respectively. If $\theta^0 \neq 0$, anomaly effects are then detected. In our framework, $g(X_{it}) = \theta^{0\prime} X_{it}$ and α_i^0 can be viewed as unknown factor loading with constant factor 1. Serlenga, Shin, and Snell (2004) consider a Wald statistic for testing the null hypothesis that $\theta^0 = 0$. Zhang (2009) adopts a similar model to check the explanatory power of firm- or asset-specific variables in cross-sections of expected returns when d_t is a vector of factors extracted from individual stock returns by PCA or from size- and book-to-market-ratio-sorted portfolio returns.

To capture the anomaly effects, we replace the constant α^0 in the traditional asset pricing models with a nonparametric function $g(X_{it})$. Then given (2.1), we can test for anomaly effects through the following hypotheses:

$$\mathbb{H}_0^{(0)} : \Pr\{g(X_{it}) = 0\} = 1 \tag{2.6}$$

versus

$$\mathbb{H}_1^{(0)}: \Pr\{g(X_{it}) = 0\} < 1. \tag{2.7}$$

That is, the absence of anomaly effects requires that $g(X_{it})$ should be zero almost surely under the null hypothesis $\mathbb{H}_0^{(0)}$, and when $g(X_{it})$ is not equal zero almost surely, we have evidence of anomaly effects under the alternative hypothesis $\mathbb{H}_1^{(0)}$.

The major difference between our test and the traditional test is that, in our test, the alternative hypothesis is tied to the given firm- or asset-specific characteristics, while in the traditional one, the alternative hypothesis is general and unspecified. Our test is more powerful than the traditional one if the mispricing, α^0, is indeed related to the given firm- or asset-specific variables, but is less powerful if the mispricing is unrelated to the given firm- or asset-specific variables. If we reject $\mathbb{H}_0^{(0)}$, it means that the additional firm- or asset-specific characteristics X_{it} can help explain the cross-section returns and the inclusion of these additional variables in the return regression is necessary. Otherwise, there is no anomaly effect caused by X_{it}. This idea is similar to Lin and Hong (2006) who apply a generalized spectral derivative test to evaluate both one factor model and Fama-French three-factor model for the Chinese capital market.

In the case of rejection, how to model the unknown function $g(X_{it})$ is also an important issue. Linearity is commonly imposed in the literature due to its simplicity and easy interpretation but may lead to misleading results when the true function g is not linear or cannot be well approximated by a linear function. So we are also interested in testing

$$\mathbb{H}_0^{(l)}: \Pr\{g(X_{it}) = X_{it}'\theta^0\} = 1 \text{ for some } \theta^0 \in \Theta \subset \mathbb{R}^d \tag{2.8}$$

versus

$$\mathbb{H}_1^{(l)}: \Pr\{g(X_{it}) \neq X_{it}'\theta\} < 1 \text{ for all } \theta \in \Theta \subset \mathbb{R}^d. \tag{2.9}$$

If $\mathbb{H}_0^{(l)}$ is rejected, a nonlinear functional form for $g(\cdot)$ would be preferred. This test can give more information about the shape of the unknown function g.

3 Estimation

In this section, we first introduce nonparametric panel data models with interactive fixed effects and the estimation method. For the purpose of illustration, we first assume that only unobserved factors are present in the model and then discuss briefly how to extend the estimation to the models with both observed and unobserved factors.

3.1 Models with unobserved factors

We first consider the following nonparametric panel data model with interactive fixed effects

$$Y_{it} = g(X_{it}) + \lambda_i^{0\prime} f_t^0 + e_{it},\ i = 1, ..., N,\ t = 1, ..., T, \tag{3.1}$$

where Y_{it} is a scalar dependent variable, g is an unknown smooth function, X_{it} is a d-dimensional vector, f_t^0 and λ_i^0 are both R-dimensional vectors and represent the unobserved common factors and factor loadings, respectively. Since $g(\cdot)$ is an unknown function, we propose to estimate $g(\cdot)$ by the method of sieves. For some excellent reviews on sieve methods, see Chen (2007, 2013). To proceed, let $p^K(x) \equiv (p_1(x), \cdots, p_K(x))'$ denote a sequence of basis functions that can approximate any square-integrable function of x very well. Then we can approximate $g(x)$ by $\beta' p^K(x)$ for some $K \times 1$ vector β under fairly weak conditions. Let $K \equiv K_{NT}$ be some integer such that $K \to \infty$ as $(N, T) \to \infty$ where $(N, T) \to \infty$ denotes that N and T pass to infinity simultaneously. We introduce the following notation: $p_{it,k} \equiv p_k(X_{it})$, $p_{it} \equiv p^K(X_{it})$, $P_i \equiv (p_{i1}, \cdots, p_{iT})'$, $P_{i\cdot,k} \equiv (p_{i1,k}, \cdots, p_{iT,k})'$, and $\mathbf{P}_k \equiv (P_{1\cdot,k}, \cdots, P_{N\cdot,k})'$. We use $\beta^0 = (\beta_1^0, ..., \beta_K^0)'$ to denote the true value of $\beta = (\beta_1, ..., \beta_K)'$ in the sieve approximation of $g(x)$ based on a linear combination of elements in $p^K(x)$. Here we suppress the dependence of p_{it}, β^0, and β on K for notational simplicity.

To estimate g, we consider the following approximating linear panel data models with interactive fixed effects:

$$Y_{it} = p_{it}'\beta^0 + \lambda_i^{0\prime} f_t^0 + u_{it} \tag{3.2}$$

where $u_{it} \equiv e_{it} + [g(X_{it}) - p_{it}'\beta^0]$ is the new error term, and $g(X_{it}) - p_{it}'\beta^0$ represents the sieve approximation error. Let $u_i \equiv (u_{i1}, \cdots, u_{iT})'$, $\mathbf{u} \equiv (u_1, \cdots, u_N)'$, $Y_i \equiv (Y_{i1}, \cdots, Y_{iT})'$, $\mathbf{Y} \equiv (Y_1, \cdots, Y_N)'$, $f \equiv (f_1, \cdots, f_T)'$ and $\lambda \equiv (\lambda_1, \cdots, \lambda_N)'$. Similarly, denote the true values of f and λ as f^0 and λ^0. In matrix notation, (3.2) can be rewritten as

$$\mathbf{Y} = \sum_{k=1}^{K} \beta_k^0 \mathbf{P}_k + \lambda^0 f^{0\prime} + \mathbf{u}. \tag{3.3}$$

Following Bai (2009), Moon and Weidner (2013a, 2013b), and Su and Zhang (2013), we estimate the model in (3.3) by the Gaussian QMLE method. That is, we obtain the estimator $(\hat{\beta}, \hat{\lambda}, \hat{f})$ of $(\beta^0, \lambda^0, f^0)$ as follows

$$(\hat{\beta}, \hat{\lambda}, \hat{f}) = \underset{(\beta,\lambda,f)}{\operatorname{argmin}} \frac{1}{NT} \operatorname{tr}\left[\left(\mathbf{Y} - \sum_{k=1}^{K} \beta_k \mathbf{P}_k - \lambda f'\right)' \left(\mathbf{Y} - \sum_{k=1}^{K} \beta_k \mathbf{P}_k - \lambda f'\right)\right], \tag{3.4}$$

where $\operatorname{tr}(\cdot)$ is the trace operator. In particular, based on the concentrated quasi-log-likelihood and the idea of principal component analysis, we can obtain $\hat{\beta}$, $\hat{\lambda}$, and $\hat{f}$ as follows:

$$\hat{\beta} = \underset{\beta \in \mathbb{R}^K}{\operatorname{argmin}} L_{NT}(\beta) \equiv \underset{\beta \in \mathbb{R}^K}{\operatorname{argmin}} \frac{1}{NT} \sum_{t=R+1}^{T} \mu_t \left[\left(\mathbf{Y} - \sum_{k=1}^{K} \beta_k \mathbf{P}_k\right)' \left(\mathbf{Y} - \sum_{k=1}^{K} \beta_k \mathbf{P}_k\right)\right], \tag{3.5}$$

$$\left[\frac{1}{NT}\left(\mathbf{Y}-\sum_{k=1}^{K}\hat{\beta}_k\mathbf{P}_k\right)'\left(\mathbf{Y}-\sum_{k=1}^{K}\hat{\beta}_k\mathbf{P}_k\right)\right]\hat{f}=\hat{f}V_{NT}, \tag{3.6}$$

and

$$\hat{\lambda}\equiv\left(\hat{\lambda}_1,\cdots,\hat{\lambda}_N\right)'=T^{-1}\left[\hat{f}'\left(Y_1-P_1\hat{\beta}\right),\cdots,\hat{f}'\left(Y_N-P_N\hat{\beta}\right)\right]', \tag{3.7}$$

where $\mu_t(A)$ denotes the t-th largest eigenvalues of a symmetric matrix A where eigenvalues of multiplicity are counted multiple times, V_{NT} is a diagonal matrix consisting of R largest eigenvalues of the matrix in the square bracket in (3.5) in decreasing order, and the estimate $\hat{f}$ of f is obtained as the corresponding first R eigenvectors. Here we use the same identification restrictions as Bai (2009): $f'f/T=I_R$ and $\lambda'\lambda$ =diagonal matrix. Multiple starting values for numerical optimization are recommended since the objective function $L_{NT}(\beta)$ is neither convex nor differentiable with respect to β. Nevertheless, we find that that it is satisfactory to apply Bai's (2009) iterative estimate of β as the starting value in the optimization in (3.5).

After obtaining $\hat{\beta}$, the sieve estimator for $g(x)$ is given by $\hat{g}(x)=p^K(x)'\hat{\beta}$. The asymptotic analysis in Su and Zhang (2013) shows that under some regularity conditions,

$$A_{NT}(x)\left[\hat{g}(x)-g(x)\right]-B_K(x)\overset{d}{\to}N(0,1)$$

as $(N,T)\to\infty$. Here $A_{NT}(x)$ and $B_K(x)$ denote the asymptotic variance and bias of $\hat{g}(x)$, respectively. Their formulae are quite complicated and omitted here, but we will present their consistent estimates below.

The asymptotic bias term $B_K(x)$ can be decomposed into two parts, corresponding to two sources of bias, denoted as $b_1(x)$ and $b_2(x)$ respectively. $b_1(x)$ is caused by cross-sectional heteroskedasticity of errors conditional on $\mathcal{D}$ – the σ-field generated by factors and factor loadings; while $b_2(x)$ is caused by serial correlation and heteroskedasticity of errors conditional on $\mathcal{D}$. In the special case where e_{it}'s are i.i.d. conditional on $\mathcal{D}$ across both i and t, the two bias terms disappear. To obtain consistent estimators of $A_{NT}(x)$ and remove these bias terms, we first obtain the residuals by

$$\hat{e}_{it}\equiv Y_{it}-\hat{g}(X_{it})-\hat{\lambda}_i'\hat{f}_t. \tag{3.8}$$

Let $\hat{\alpha}_{ij}\equiv\hat{\lambda}_i'(\hat{\lambda}'\hat{\lambda}/N)^{-1}\hat{\lambda}_j$, $\hat{\eta}_{ts}\equiv\hat{f}_t'(\hat{f}'\hat{f}/T)^{-1}\hat{f}_s$, and $\hat{Z}_{it}\equiv p_{it}-\frac{1}{N}\sum_{j=1}^{N}\hat{\alpha}_{ij}p_{jt}-\frac{1}{T}\sum_{s=1}^{T}\hat{\eta}_{ts}p_{is}+\frac{1}{NT}\sum_{j=1}^{N}\sum_{s=1}^{T}\hat{\alpha}_{ij}\hat{\eta}_{ts}p_{js}$. Define

$$\hat{W}_{NT}\equiv\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\hat{Z}_{it}\hat{Z}_{it}',\ \hat{\Omega}_{NT}\equiv\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\hat{Z}_{it}\hat{Z}_{it}'\hat{e}_{it}^2, \tag{3.9}$$

$$\hat{V}_K(x)\equiv p^K(x)'\hat{W}_{NT}^{-1}\hat{\Omega}_{NT}\hat{W}_{NT}^{-1}p^K(x),\ \text{and } \hat{A}_{NT}(x)\equiv\sqrt{NT/\hat{V}_K(x)}. \tag{3.10}$$

Let $\hat{e}_i=(\hat{e}_{i1},...,\hat{e}_{iT})'$, $\hat{\mathbf{e}}=(\hat{e}_1,...,\hat{e}_N)'$, $\hat{\Phi}=\hat{\lambda}(\hat{\lambda}'\hat{\lambda})^{-1}(\hat{f}'\hat{f})^{-1}\hat{f}$, $A^{\mathrm{D}}=\mathrm{diag}(A_{11},...,A_{nn})$ for any $n\times n$ matrix A with (i,j)th element A_{ij}, and $M_B=I_n-B\left(B'B\right)^{-1}B'$ for any $n\times m$ matrix B, where I_n is an $n\times n$ identify matrix. Define $K\times 1$ vectors $\hat{b}_1$ and $\hat{b}_2$ whose kth elements are respectively given by

$$\hat{b}_{1,k}\equiv\frac{1}{T}\mathrm{tr}\left[\left(\hat{\mathbf{e}}\hat{\mathbf{e}}'\right)^{\mathrm{D}}M_{\hat{\lambda}}\mathbf{P}_k\hat{\Phi}\right]\text{ and }\hat{b}_{2,k}\equiv\frac{1}{N}\mathrm{tr}\left[\left(\hat{\mathbf{e}}'\hat{\mathbf{e}}\right)^{\mathrm{D}}M_{\hat{f}}\mathbf{P}_k'\hat{\Phi}'\right]. \tag{3.11}$$

Then $\hat{b}_l(x)=-\hat{A}_{NT}p^K(x)'\hat{W}_{NT}^{-1}\hat{b}_l/N$ estimates for the bias $b_l(x)$ for $l=1,2$. The bias-corrected estimator of β is given by

$$\hat{\beta}_{bc}\equiv\hat{\beta}+\hat{W}_{NT}^{-1}(N^{-1}\hat{b}_1+T^{-1}\hat{b}_2). \tag{3.12}$$

Let $\hat{B}_K(x)=-\hat{A}_{NT}p^K(x)'\hat{W}_{NT}^{-1}(N^{-1}\hat{b}_1+T^{-1}\hat{b}_2)=\hat{b}_1(x)+\hat{b}_2(x)$. The bias-corrected estimator of $g(x)$ is given by

$$\hat{g}_{bc}(x)\equiv p^K(x)'\hat{\beta}_{bc}=\hat{g}(x)-\hat{A}_{NT}^{-1}\hat{B}_K(x). \tag{3.13}$$

Under some regularity conditions, Su and Zhang (2013) show that $\hat{A}_{NT}(x)[\hat{g}_{bc}(x)-g(x)]\xrightarrow{d}N(0,1)$ as $(N,T)\to\infty$.

To construct our test statistic for testing the null hypothesis in (2.8), we also need to estimate the model under $\mathbb{H}_0^{(l)}$. In this case, $g(x)=x'\theta^0$. By replacing $\{p_{it}\}$ with $\{X_{it}\}$ in the sieve estimation, we can obtain the linear estimators $(\hat{\theta},\hat{\lambda},\hat{f})$ of $\left(\theta^0,\lambda^0,f^0\right)$. Similarly, we can define a bias-corrected estimate for θ^0 : $\hat{\theta}_{bc}\equiv\hat{\theta}+\hat{D}_{NT}^{-1}(N^{-1}\hat{b}_{l,1}+T^{-1}\hat{b}_{l,2})$, where $\hat{D}_{NT}$, $\hat{b}_{l,1}$ and $\hat{b}_{l,2}$ are defined similarly to $\hat{W}_{NT}$, $\hat{b}_1$ and $\hat{b}_2$, respectively, but with $\{p_{it}\}$ being replaced by $\{X_{it}\}$. Su, Jin, and Zhang (2013) show that $\sqrt{NT}(\hat{\theta}_{bc}-\theta^0)\xrightarrow{d}N\left(0,D_0^{-1}\Omega_eD_0^{-1}\right)$ under some conditions, where $D_0\equiv\mathrm{plim}_{(N,T)\to\infty}D_{NT}$, $D_{NT}\equiv\frac{1}{NT}\sum_{i=1}^N\sum_{t=1}^T\tilde{X}_{it}\tilde{X}_{it}'$, $\Omega_e\equiv\mathrm{plim}_{(N,T)\to\infty}\frac{1}{NT}\sum_{i=1}^N\sum_{t=1}^T\tilde{X}_{it}\tilde{X}_{it}'e_{it}^2$, $\tilde{X}_{it}$ is the tth row of $\tilde{X}_i\equiv M_{f^0}X_i-\frac{1}{N}\sum_{j=1}^N\alpha_{ij}M_{f^0}X_j$, and $\alpha_{ij}\equiv\lambda_i^{0\prime}(\lambda^{0\prime}\lambda^0/N)^{-1}\lambda_j^0$.

3.2 Models with both observed and unobserved factors

We now consider nonparametric panel data models with both observed and unobserved factors:

$$Y_{it}=g(X_{it})+\Gamma_i^{0\prime}d_t+\lambda_i^{0\prime}f_t^0+e_{it}, \tag{3.14}$$

where the notation on the right hand side follows from that in equation (2.1). To estimate g, we consider the following approximating linear panel data models with interactive fixed effects:

$$Y_{it}=p_{it}'\beta^0+\Gamma_i^{0\prime}d_t+\lambda_i^{0\prime}f_t^0+u_{it}, \tag{3.15}$$

where $u_{it}=e_{it}+\left[g(X_{it})-p_{it}'\beta^0\right]$. In vector notation, (3.15) can be rewritten as

$$Y_i=P_i\beta^0+D\Gamma_i^0+f^0\lambda_i^0+u_i \tag{3.16}$$

where $D=(d_1,...,d_T)'$ and $u_i=(u_{i1},...,u_{iT})'$. Denote $\tilde{A}=M_DA$ for $A=Y_i$, P_i, f^0, or u_i, where $M_D=I_T-D(D'D)^{-1}D'$. Multiplying M_D on both sides of (3.16) yields

$$\tilde{Y}_i=\tilde{P}_i\beta^0+\tilde{f}^0\lambda_i^0+\tilde{u}_i. \tag{3.17}$$

Now we obtain a panel data model only with unobserved factors $\tilde{f}^0$. We can follow the method introduced before to estimate β^0, $\tilde{f}^0$ and λ^0. For the purpose of identification, it is reasonable to impose an additional restriction that the observed factors are orthogonal to the unobserved factors. That is, we assume that the space spanned by the columns of D is orthogonal to the space spanned by the columns of f^0, i.e., $D'f^0 = 0$. With this identification restriction, we have: $\tilde{f}^0 = M_D f^0 = [I_T - D(D'D)^{-1} D']f^0 = f^0$. Consequently, we can estimate the model in (3.17) via the Gaussian QMLE method with $\tilde{f}^0$ being replaced by f^0. Specifically, we obtain the estimator $(\hat{\beta}, \hat{\lambda}, \hat{f})$ of $(\beta^0, \lambda^0, f^0)$ as follows

$$(\hat{\beta}, \hat{\lambda}, \hat{f}) = \underset{(\beta,\lambda,f)}{\arg\min} \frac{1}{NT} \sum_{i=1}^{N} \left(\tilde{Y}_i - \tilde{P}_i\beta - f\lambda_i\right)' \left(\tilde{Y}_i - \tilde{P}_i\beta - f\lambda_i\right). \tag{3.18}$$

Once we get $\hat{\beta}, \hat{\lambda}$, and $\hat{f}$, the factor loadings Γ_i^0's can be estimated by

$$\hat{\Gamma}_i = (D'D)^{-1} D' \left(Y_i - P_i\hat{\beta} - \hat{f}\hat{\lambda}_i\right).$$

We can estimate e_{it} by $\hat{e}_{it} = Y_{it} - p'_{it}\hat{\beta} - \hat{\Gamma}'_i d_t - \hat{f}'_t\hat{\lambda}_i$. Let $\tilde{p}_{it} = p_{it} - \frac{1}{T}\sum_{s=1}^{T} d'_s (D'D/T)^{-1} d_t p_{is}$ and $\tilde{Z}_{it} \equiv \tilde{p}_{it} - \frac{1}{N}\sum_{j=1}^{N} \hat{\alpha}_{ij}\tilde{p}_{jt} - \frac{1}{T}\sum_{s=1}^{T} \hat{\eta}_{ts}\tilde{p}_{is} + \frac{1}{NT}\sum_{j=1}^{N}\sum_{s=1}^{T} \hat{\alpha}_{ij}\hat{\eta}_{ts}\tilde{p}_{js}$, where $\hat{\eta}_{ts}$ and $\hat{\alpha}_{ij}$ are defined as before. Define $\tilde{W}_{NT}$, $\tilde{\Omega}_{NT}$, $\tilde{V}_K(x)$ and $\tilde{A}_{NT}(x)$ analogously to $\hat{W}_{NT}$, $\hat{\Omega}_{NT}$, $\hat{V}_K(x)$ and $\hat{A}_{NT}(x)$, respectively, but with $\{\hat{Z}_{it}\}$ being replaced by $\{\tilde{Z}_{it}\}$. Construct $\tilde{b}_1$ and $\tilde{b}_2$ as the estimates of b_1 and b_2 by using $\hat{f}$, $\hat{\lambda}$ and $\{\hat{e}_{it}\}$ and replacing $\{p_{it}\}$ with $\{\tilde{p}_{it}\}$ in (3.11) and (3.12). Then we can define a bias-corrected estimate of $g(x)$ by $\hat{g}_{bc}(x) = p^K(x)'\hat{\beta}_{bc}$, where $\hat{\beta}_{bc} = \hat{\beta} + \tilde{W}_{NT}^{-1}(N^{-1}\tilde{b}_1 + T^{-1}\tilde{b}_2)$. Following Su and Zhang (2013), one can readily establish the claim that $\tilde{A}_{NT}(x)[\hat{g}_{bc}(x) - g(x)] \xrightarrow{d} N(0,1)$ under some regular conditions.

4 Testing for the anomaly effects and functional forms

In this section, we consider specification tests for the commonly used functional forms in panel data models with interactive fixed effects. For the model in (3.1) or (3.14), we are interested in testing the null hypothesis

$$\mathbb{H}_0^{(l)} : \Pr\left[g(X_{it}) = X'_{it}\theta^0\right] = 1 \text{ for some } \theta^0 \in \Theta, \tag{4.1}$$

where Θ is a compact subset of $\mathbb{R}^d$. The alternative hypothesis is $\mathbb{H}_1^{(l)} : \Pr[g(X_{it}) = X'_{it}\theta] < 1$ for all $\theta \in \Theta$. The test statistic is based on the comparison between the linear estimator under the null hypothesis and the sieve estimator under the alternative; see Su and Zhang (2013) for details.

In empirical applications in economics and finance, people are also interested in testing the following null hypothesis

$$\mathbb{H}_0^{(0)} : \Pr\left[g(X_{it}) = 0\right] = 1. \tag{4.2}$$

The alternative is $\mathbb{H}_1^{(0)}$: $\Pr[g(X_{it}) = 0] < 1$. Obviously, the null hypothesis in (4.2) can be regarded as a special case of that in (4.1) by taking $\theta^0 = 0$. As mentioned above, $\mathbb{H}_0^{(0)}$ indicates the absence of anomaly effects.

4.1 The test statistic

We first introduce Su and Zhang's (2013) test statistic that is based on the weighted L_2-distance between two estimators of $g(\cdot)$, i.e., the linear and sieve estimators. Intuitively, both estimators are consistent under the null hypothesis of linearity whereas only the sieve estimator is consistent under the alternative hypothesis. So if there is any deviation from the null, the L_2-distance between the two estimators will signal it out asymptotically. This motivates the following test statistic

$$\Gamma_{s,NT} \equiv \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} \left[\hat{g}_{bc}(X_{it}) - \hat{g}_{(l)}(X_{it}) \right]^2 w(X_{it}),$$

where $\hat{g}_{(l)}(x) = x'\hat{\theta}$, $\hat{\theta}$ is Moon and Weidner's (2013a, 2013b) linear estimator of the coefficient θ^0 under $\mathbb{H}_0^{(l)}$ (bias correction is preferred and used below but not required in theory), and $w(x)$ is a user-specified nonnegative weighting function. Similar test statistics have been proposed in various other contexts in the literature; see, e.g., Härdle and Mammen (1993), Hong and White (1995), and Su and Lu (2013). After being appropriately centered and scaled, $\Gamma_{s,NT}$ is asymptotically normally distributed under the null hypothesis of linearity.

Define

$$\hat{\mathbb{B}}_{s,NT} \equiv \operatorname{tr}\left(\hat{W}_{NT}^{-1} Q_{NT} \hat{W}_{NT}^{-1} \hat{\Omega}_{NT} \right) \text{ and } \hat{\mathbb{V}}_{s,NT} \equiv 2\operatorname{tr}\left(\hat{W}_{NT}^{-1} Q_{NT} \hat{W}_{NT}^{-1} \hat{\Omega}_{NT} \hat{W}_{NT}^{-1} Q_{NT} \hat{W}_{NT}^{-1} \hat{\Omega}_{NT} \right).$$

Here the definitions of Q_{NT}, $\hat{W}_{NT}$, and $\hat{\Omega}_{NT}$ depend on whether the observed factor d_t is present or not (see model (3.1) or (3.14)). Let $w_{it} \equiv w(X_{it})$. When d_t is absent as in model (3.1), $Q_{NT} = \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} p_{it} p_{it}' w_{it}$, and $\hat{W}_{NT}$ and $\hat{\Omega}_{NT}$ are defined as in Section 3.1. When d_t is present in model (3.14), $Q_{NT} = \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} \tilde{p}_{it} \tilde{p}_{it}' w_{it}$, $\hat{W}_{NT} \equiv \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} \tilde{Z}_{it} \tilde{Z}_{it}'$, and $\hat{\Omega}_{NT} \equiv \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} \tilde{Z}_{it} \tilde{Z}_{it}' \hat{e}_{it}^2$, where $\tilde{Z}_{it}$ and $\tilde{p}_{it}$ are as defined in Section 3.2. Then we define a feasible test statistic:

$$\hat{J}_{s,NT} \equiv \left(NT\, \Gamma_{s,NT} - \hat{\mathbb{B}}_{s,NT} \right) / \sqrt{\hat{\mathbb{V}}_{s,NT}}. \tag{4.3}$$

Su and Zhang (2013) show that $\hat{J}_{s,NT} \xrightarrow{d} N(0,1)$ under $\mathbb{H}_0^{(l)}$.

We can also modify our test statistic to test for the null hypothesis $\mathbb{H}_0^{(0)}$ of no anomaly effects. Define

$$\Gamma_{s0,NT} \equiv \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} \left[\hat{g}_{bc}(X_{it}) - 0 \right]^2 w(X_{it})$$

which compares the distance between the sieve estimate of $g(\cdot)$ and 0. It is easy to show that under $\mathbb{H}_0^{(0)}$,

$$\hat{J}_{s0,NT} \equiv \left(NT\, \Gamma_{s0,NT} - \hat{\mathbb{B}}_{s,NT} \right) / \sqrt{\hat{\mathbb{V}}_{s,NT}} \xrightarrow{d} N(0,1).$$

4.2 A bootstrap version of the test

Despite the fact that $\hat{J}_{s,NT}$ and $\hat{J}_{s0,NT}$ are asymptotically $N(0,1)$ under the null, it is not recommended to rely on the asymptotic normal critical values to make statistical inference in finite samples because of the nonparametric nature of the tests. In addition, even though the slow convergence rates of the factors and factor loadings estimates do not affect the asymptotic normal distributions of the test statistics, they tend to have adverse effects in finite samples (see, Su and Chen, 2013). As a result, tests based on standard normal critical values tend to suffer severe size distortions in finite samples. For this reason, Su and Zhang (2013) propose a fixed-regressor wild bootstrap method for the test of $\mathbb{H}_0^{(l)}$ and demonstrate their superb finite sample performance. The bootstrap procedure goes as follows:

1. Under $\mathbb{H}_0^{(l)}$, obtain the linear estimators $\hat{\theta}$, $\hat{f}_t^{(l)}$, $\hat{\lambda}_i^{(l)}$, and $\hat{e}_{it}^{(l)}$, where the superscript (l) denotes estimates under the null hypothesis of linearity; under $\mathbb{H}_1^{(l)}$, obtain the bias-corrected sieve estimators: $\hat{\beta}_{bc}$, $\hat{f}_t$, $\hat{\lambda}_i$, and $\hat{e}_{it}$. Calculate the test statistic $\hat{J}_{s,NT}$ based on $\hat{g}_{bc}(X_{it}) = \hat{\beta}_{bc}' p^K(X_{it})$, $\hat{\theta}' X_{it}$, $\hat{\lambda}_i$, $\hat{f}_t$, and $\hat{e}_{it}$.

2. For $i = 1, ..., N$, obtain the wild bootstrap errors $\{e_{it}^*\}_{t=1}^T$ as follows: $e_{it}^* = \nu_{it}\hat{e}_{it}^{(l)}$ where ν_{it} are i.i.d. $N(0,1)$. Then generate the bootstrap analogue Y_{it}^* of Y_{it} by holding $(X_{it}, \hat{f}_t^{(l)}, \hat{\lambda}_i^{(l)})$ as fixed: $Y_{it}^* = X_{it}'\hat{\theta} + \hat{\lambda}_i^{(l)\prime} \hat{f}_t^{(l)} + e_{it}^*$ for $i = 1, ..., N$ and $t = 1, ..., T$.

3. Given the bootstrap resample $\{Y_{it}^*, X_{it}\}$, obtain the sieve QMLEs $\hat{g}_{bc}^*(X_{it})$, $\hat{\lambda}_i^*$, $\hat{f}_t^*$ and $\hat{e}_{it}^*$, and the linear estimators $\hat{\theta}^*$, $\hat{\lambda}_i^{(l)*}$, $\hat{f}_t^{(l)*}$ and $\hat{e}_{it}^{(l)*}$. Calculate the bootstrap test statistic $\hat{J}_{NT}^*$ based on $\hat{g}_{bc}^*(X_{it})$, $X_{it}'\hat{\theta}^*$, $\hat{f}_t^*$, $\hat{\lambda}_i^*$, and $\hat{e}_{it}^*$.

4. Repeat Steps 2-3 for B times and index the bootstrap statistics as $\{\hat{J}_{NT,b}^*\}_{b=1}^B$. Calculate the bootstrap p-value: $p^* = B^{-1}\sum_{b=1}^B \mathbf{1}\{\hat{J}_{NT,b}^* \geq \hat{J}_{NT}\}$ where $\mathbf{1}\{\cdot\}$ is the usual indicator function.

It is straightforward to implement the above bootstrap procedure. Note that we impose the null hypothesis of linearity in Step 2. Since the regressors are treated as fixed, there is no dynamic structure in the bootstrap world. With a minor modification, the above bootstrap procedure can be applied to the test of $\mathbb{H}_0^{(0)}$ versus $\mathbb{H}_1^{(0)}$. The difference is that here we need to estimate a pure panel factor model instead of a linear panel data model with a factor structure in Steps 1 and 3, and generate the bootstrap resamples using the residuals from the pure panel factor model.

5 Empirical application

In this section we apply the specification tests to test for anomaly effects in empirical asset pricing models. We first discuss the data and implementation, and then report the test and estimation results.

5.1 Data

We collect monthly data on the average value weighted excess returns and size (SZ), and annual data on book-to-market ratio (BM) for 100 constructed portfolios for the period from June 1973 to December 2009 from Kenneth French's website.[1] A total of N =96 portfolios are available for the selected sample period. We collect the monthly Fama/French's three factors, i.e., SMB, HML, and market excess return (MKT), and monthly Momentum factor (Mom) from Kenneth French's data library. In addition to the above four factors, we also download monthly data on the market-wide liquidity factor (Liq, the level of aggregate liquidity) from Pástor's website.[2] To remove the outliers of the return data, we truncate the data using 97.5% percentile of the original data as upper bound and 2.5% quantile as a lower bound. In addition to the full sample 1973/07-2009/12 (T = 438), we also consider three subsamples: 1973/07–1983/12 (T = 126), 1984/01–1996/12 (T = 156), and 1997/01–2009/12 (T = 156). The first subsample is the period between the end of oil shock in 1972 and the starting of "Great Moderation" in 1984. The last subsample corresponds to the post Asian financial crisis period. The descriptive statistics about our data set are presented in Table 1. From Table 1 we can observe a large variation in some of the financial variables, e.g., Mom and Liq, across the three subsample periods. For the Mom, the standard deviations vary from 3.1698 in the second subsample to 6.2755 in the third subsample, and for the Liq, the standard deviations vary from 3.6576 in the second subsample to 8.3147 in the third subsample. The two characteristic variables, SZ and BM, are quite stable, with standard deviations around 1.5127 to 1.8846 for SZ and 0.5792 to 0.8792 for BM.

5.2 Implementation

5.2.1 Models

We consider different tests for five models: (1) one-factor model (CAPM), (2) Fama-French three-factor models (FF), (3) 4-factor models (4F, Fama-French three factors plus momentum factor), (4) 5-factor models (5F, Fama-French three factors plus momentum factor and liquidity factor), and (5) pure unobserved factor models (PUF).

5.2.2 Tests

To test the anomaly effects in the portfolios returns, we consider three different tests. The first one is the sieve-based test with different numbers of knots used in the construction of the cubic B-spline sieve bases (see Section 5.2.5 below), the second one is the parametric Wald test based on the linear estimator $\hat{\theta}_{bc}$ introduced above where we let the data determine the number of unobserved factors, and the third one is the parametric Wald test based on a linear specification of $g(\cdot)$ without including any

[1] Website: mba.tuck.dartmouth.edu/pages/faculty/ken.french/Data_Library/det_100_port_sz.html.
[2] Website: http://faculty.chicagobooth.edu/lubos.pastor/research/liq_data_1962_2012.txt.

Table 1: Descriptive statistics for variables (96 portfolios) and factors

Variable or factor	Period	Mean	Median	Std deviation	Max	Min
Excess return (×100)	1973/07-2009/12	0.6882	0.8300	6.1345	94.9400	-42.6100
	1973/07-1983/12	0.7679	0.3000	6.7217	37.0200	-31.8900
	1984/01-1996/12	0.8945	1.0700	4.9472	26.3200	-26.3100
	1997/01-2009/12	0.5903	0.8950	6.8470	94.9400	-42.6100
SZ	1973/07-2009/12	6.1639	6.1742	1.8846	11.7025	1.3481
	1973/07-1983/12	4.8327	4.7612	1.5127	8.8091	1.3481
	1984/01-1996/12	5.9529	5.9712	1.6889	10.4153	1.9544
	1997/01-2009/12	7.1509	7.1231	1.7166	11.7025	3.0516
BM	1973/07-2009/12	0.8995	0.7200	0.7007	6.5800	0.0800
	1973/07-1983/12	1.176	0.9550	0.8792	6.5800	0.1300
	1984/01-1996/12	0.9030	0.7800	0.5792	4.6400	0.1100
	1997/01-2009/12	0.6877	0.5400	0.5834	5.2300	0.0800
MKT (×100)	1973/07-2009/12	0.4228	0.7750	4.5288	16.1000	-17.2300
	1973/07-1983/12	0.3454	0.0350	4.7622	16.1000	-12.7500
	1984/01-1996/12	0.8004	1.0500	3.8719	11.3000	-12.9000
	1997/01-2009/12	0.2444	1.1350	4.9207	10.1900	-17.2300
SMB (×100)	1973/07-2009/12	0.2258	0.0700	3.2334	22.0000	-16.3900
	1973/07-1983/12	0.4418	0.2800	3.3846	11.0100	-9.9000
	1984/01-1996/12	0.0639	0.0000	2.3337	8.4700	-6.6400
	1997/01-2009/12	0.3144	0.0100	3.9248	22.0000	-16.3900
HML (×100)	1973/07-2009/12	0.4549	0.4700	3.1251	13.8400	-12.6000
	1973/07-1983/12	0.5204	0.7300	2.7101	8.6000	-9.7600
	1984/01-1996/12	0.3437	0.3500	2.6262	7.6100	-8.4400
	1997/01-2009/12	0.3598	0.4050	3.7557	13.8400	-12.6000
Mom (×100)	1973/07-2009/12	0.7884	0.9350	4.6853	18.3900	-34.7200
	1973/07-1983/12	0.8496	0.8400	3.8199	10.2600	-13.8000
	1984/01-1996/12	1.0104	1.2850	3.1698	15.2400	-9.5800
	1997/01-2009/12	0.4325	0.8500	6.2755	18.3900	-34.7200
Liq (×100)	1973/07-2009/12	-3.2304	-2.3204	6.3761	20.1015	-33.3629
	1973/07-1983/12	-5.1489	-4.4602	5.5665	9.6240	-30.0164
	1984/01-1996/12	-1.5002	-1.3889	3.6576	8.0346	-15.7862
	1997/01-2009/12	-3.7786	-3.4126	8.3147	20.1015	-33.3629

unobserved factors. The last one is similar to the test in Zhang (2009) who studies anomaly effects by using linear regression models with constructed factors. For the latter two tests, the Wald statistics are used to test where $\theta^0 = 0$ when $g(\cdot)$ is linearly specified: $g(x) = x'\theta^0$.

For the specification test of linear functional forms which is especially useful when anomaly effects are detected, we only consider the sieve-based test.

5.2.3 Determination of the number of sieve terms

In sieve estimation, we have to choose the number of sieve terms. There are three common methods to select the number of approximating terms given sieve bases. The first one is a data-driven method which selects the "optimal" number of sieve terms based on least squares cross-validation (CV), generalized CV (GCV), or some information criteria (e.g., AIC, BIC). This is motivated from the apparent connection of the method of sieves with the parametric method. For example, Andrews (1991) establishes the asymptotic optimality of CV as a method to select series terms for nonparametric least square regression with heteroskedastic errors. The second method is to apply *lasso* (least absolute shrinkage and selection operator) to select the significant terms and estimate the model simultaneously in nonparametric least square regression; see Belloni and Chernozhukov (2013). The third method is to consider a candidate set of numbers of sieve terms to evaluate the sensitivity of the estimation and testing to different choices of numbers of sieve terms.

To the best of our knowledge, there is no theoretical work on the asymptotic properties of CV, GCV, AIC, BIC, or lasso for the nonparametric panel data models with factor structural errors. For this reason, we adopt the third approach in our application.

5.2.4 Choice of the number of factors

In practice, we also need to choose the number of factors. Bai and Ng (2002) consider determining the number of factors in pure unobserved factor models without any regressors. As Su, Jin, and Zhang (2013) remark, it is easy to extend their theory to linear panel data models with interactive fixed effects.[3] Under either the null of linearity ($\mathbb{H}_0^{(l)}$) or the null of no anomaly effect ($\mathbb{H}_0^{(0)}$), we have a linear panel data model with interactive fixed effects. Following Bai and Ng (2002), we can adopt the

[3] See Lu and Su (2014) for the performance of Bai and Ng's (2002) information criteria in linear panel data models with interactive fixed effects.

following recommended information criteria to choose the number of factors:

$$\begin{aligned} PC_{p1}(R) &= V\left(R, \hat{f}^R\right) + R\hat{\sigma}^2 \left(\frac{N+T}{NT}\right) \ln\left(\frac{NT}{N+T}\right), \\ PC_{p2}(R) &= V\left(R, \hat{f}^R\right) + R\hat{\sigma}^2 \left(\frac{N+T}{NT}\right) \ln C_{NT}^2, \\ IC_{p1}(R) &= \ln\left(V\left(R, \hat{f}^R\right)\right) + R\left(\frac{N+T}{NT}\right) \ln\left(\frac{NT}{N+T}\right), \\ IC_{p2}(R) &= \ln\left(V\left(R, \hat{f}^R\right)\right) + R\left(\frac{N+T}{NT}\right) \ln C_{NT}^2, \end{aligned}$$

where $C_{NT}^2 = \min\{N, T\}$, $V(R, \hat{f}^R) = \frac{1}{NT}\sum_{i=1}^N \sum_{t=1}^T \left(\hat{e}_{it}^R\right)^2$, $\hat{e}_{it}^R$ is the residual and $\hat{f}^R$ is estimate of the factor matrix f when R factors are used and the model is estimated under the null of linearity ($\mathbb{H}_0^{(l)}$), and $\hat{\sigma}^2$ is a consistent estimate of $\frac{1}{NT}\sum_{i=1}^N \sum_{t=1}^T E\left(e_{it}^2\right)$ under the null and can be replaced by $V(R_{\max}, \hat{f}^{R_{\max}})$ in applications. Following Bai and Ng (2002) we consider $R_{\max} = 8$, 10, and 15. It is clear that $PC_{p1}(R)$ and $PC_{p2}(R)$ depend on the choice of $R_{\max}$ through $\hat{\sigma}^2$ and that different criteria may yield different choices of optimal numbers of factors. Therefore, we choose the number of factors that have the majority recommendation from these four criteria and three choices of $R_{\max}$ in the estimation under the null. Where there is a tie, we use the larger number of factors. Su, Jin, and Zhang (2013) use a similar method to choose the number of factors in their panel study of economic growth with interactive fixed effects.

5.2.5 Other details

Throughout the estimation and testing, we impose additivity assumption on the nonparametric function such that $g(SZ_{it}, BM_{it}) = g_1(SZ_{it}) + g_2(BM_{it})$ for simplicity. Note that almost all the empirical works in the literature impose additivity on testing or modeling anomaly effects. The results without imposing the additivity assumption are similar and available upon request.

In the sieve-based test, we use the cubic B-spline as the sieve basis. To construct the cubic B-spline bases for either $g_1(\cdot)$ or $g_2(\cdot)$, we need to determine the number of knots and use the same number (q) of knots in the construction for each additive component. As mentioned in Section 5.2.3, we will choose $q = 3, 4, ..., 8$ and investigate the sensitivity of our estimation and testing to the choice of q. Details on the construction of cubic B-spline can be found in Su and Zhang (2013). In addition, we need to choose a weight function $w(\cdot)$. We set $w(X_{it}) = \Pi_{l=1}^2 \mathbf{1}\{q_{l,0.025} \leq X_{l,it} \leq q_{l,0.975}\}$ where $X_{l,it}$ denotes the lth element in $X_{it} = (SZ_{it}, BM_{it})'$, and $q_{l,\tau}$ denotes the empirical τth quantile of $X_{l,it}$.[4]

We use 1000 bootstrap resamples to obtain the bootstrap p-values for the sieve-based tests for either the anomaly effects or linear functional forms, and the corresponding Wald test for the anomaly effects based on linear regression models.

[4] This weight function can be replaced by the constant 1 since we truncate the 2.5% tail observations before implementing the sieve-based and kernel-based tests.

Table 2: Bootstrap p-values for different tests of anomaly effects in various asset pricing models

	$\mathbb{H}_0^{(0)}$: $\Pr[g(X_{it})=0]=1$								
	R	$J_{l,0}$	J_l	$J_{s,3}$	$J_{s,4}$	$J_{s,5}$	$J_{s,6}$	$J_{s,7}$	$J_{s,8}$
1973/07-2009/12 (full sample)									
PUF	5	–	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
CAPM	4	<0.001	0.0138	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
FF	3	0.199	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
4F	3	<0.001	<0.001	0.001	<0.001	<0.001	<0.001	<0.001	<0.001
5F	3	<0.001	<0.001	0.001	<0.001	<0.001	<0.001	<0.001	0.001
1973/07-1983/12									
PUF	4	–	0.552	0.272	0.215	0.207	0.209	0.186	0.163
CAPM	4	<0.001	<0.001	0.002	0.002	0.001	<0.001	<0.001	<0.001
FF	2	0.041	0.123	0.187	0.180	0.200	0.191	0.151	0.143
4F	2	0.294	0.430	0.674	0.712	0.656	0.602	0.469	0.437
5F	2	0.047	0.556	0.441	0.484	0.474	0.442	0.354	0.352
1984/01-1996/12									
PUF	4	–	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
CAPM	3	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
FF	1	0.715	0.1197	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
4F	1	0.240	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
5F	1	0.130	0.0196	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
1997/01-2009/12									
PUF	7	–	<0.001	<0.001	0.015	0.006	0.004	0.002	0.007
CAPM	7	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
FF	4	0.0893	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
4F	4	<0.001	0.0496	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
5F	4	<0.001	0.0901	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001

Note: $J_{l,0}$, J_l, and $J_{s,q}$ (q=3, 4,...,8) denote the Wald test based on the linear model without unobserved factors, the Wald test based on the linear model with unobserved factors, and the sieve test with q knots, respectively.

5.3 Test results

5.3.1 Test for anomaly effects

Table 2 reports the p-values for various tests of anomaly effects in different factor pricing models. We consider the tests for both the full sample and the three subsamples. For the full sample, we can see that different numbers of unobserved factors are chosen for different factor pricing models based on the procedure discussed above. We choose 5, 4, 3, 3, and 3 unobserved factors in PUF, CAPM, FF, 4-factor, and 5-factor models, respectively. The inclusion of MKT in CAPM excludes one unobserved factor, which implies that MKT is useful to explain the returns. However, the inclusion of SMB, HML, and MKT in FF only excludes two unobserved factors. One possible explanation is that there may exist some competition between SZ and SMB, or between BM and HML. For example, if SZ has played the role of SMB in explaining the returns, then the inclusion of SMB may not result in the exclusion of any latent factors. Further, the inclusion of one or two additional factors in the 4-factor or 5-factor models does not exclude any additional unobserved factor either, which means that the unexplained factors in FF cannot be explained by Mom or Mom and Liq together. In other words, the unexplained factors are uncorrelated with Mom and Liq. Note that J_l is the Wald test for testing $\theta^0 = 0$ based on the linear estimate $\hat{\theta}_{bc}$ when one assumes that $g(x) = \theta^{0\prime} x$ in the panel model with unobserved factors. We can see that the p-values for the linear-regression-based Wald tests are all smaller than 0.05 for all five factor pricing models. Similarly, the p-values for the sieve-based tests of anomaly effects are all smaller than 0.01 for all choices of numbers of sieve terms across all five factor pricing models under investigation. So we can readily reject the null of no anomaly effects in both the linear models and nonparametric models at the 5% significance level and the two types of tests yield the same conclusion.

Table 2 suggests that the results for the three subsamples are mixed. For the first subsample, both the linear-regression-based Wald test and the sieve-based test can only reject the null of no anomaly effects at the 1% significance level in CAPM, which seems to be hard to explain. For the second subsample, except J_l in FF, all the tests can reject the null of no anomaly effects at the 5% significance level. For the third subsample, we can reject the null of no anomaly effects at the 5% significance level for almost all tests in all factor pricing models: an exception occurs for the J_l test in the 5-factor model, where we can reject the null only at the 10% significance level. Note that in this last subsample, we tend to choose a larger number of unobserved factors than in the other two subsamples. One possible explanation is that a structural change may occur across the three subsamples or within this last subsample (say, it may be caused by the shock of 9/11, or aggressively low federal funds rate or steadily increasing oil prices in this period). To the best of our knowledge, there is no test that can be used to test for the structural change in our model formally. The theoretical investigation of such an issue is certainly important and we leave it for future research.

In sum, when the linear-regression-based Wald test (J_l) reveals the presence of anomaly effects, our nonparametric sieve test also does so. But there are cases where the J_l test fails to detect the anomaly effects but our nonparametric sieve test detects. This implies that our nonparametric test is more powerful than the Wald test in detecting anomaly effects. On the other hand, our nonparametric test rejects the null of no anomalies in the full sample and the last two subsamples for all factor pricing models. This implies that the asset-specific characteristics can affect the returns even if their corresponding factors are included in the regression. In other words, our findings suggest that SMB and HML only explain a proportion of the effects of SZ and BM on returns but not all, and we do observe the evidence of anomaly effects caused by SZ and BM.

Note that Table 2 also reports the $J_{l,0}$ test results based on the linear regression model without unobserved factors, which was frequently done in the financial literature before the wide use of latent factor models. From Table 2, we can see that the p-values for the $J_{l,0}$ test in the full sample are smaller than 1% in all factor pricing models except for FF, which is largely consistent with the results based on either the J_l test or the sieve test. But the $J_{l,0}$ test results for the three subsamples can be quite different from those based on the J_l test or the sieve test. It is worth mentioning that in the presence of unobserved factors that are correlated with the regressors, the $J_{l,0}$ test is asymptotically invalid so that its test results are not trustworthy. Table 2 suggests the importance of inclusion of some unobserved factors in the factor pricing model when conducting a test for the presence of anomaly effects.

Comparison with existing results Our test for the null of no anomaly effects are closely related to Brennan, Chordia, and Subrahmanyam (1998), Zhang (2009), and Chordia, Goyal, Shanken (2012). Generally speaking, our results are largely consistent with the findings in the first and third papers; both of which detect significant anomaly effects even if both groups of authors only consider linear models. Using a variant of the two-pass test, Brennan, Chordia, and Subrahmanyam (1998) study the relationship between the firm-specific variables and two sets of factors, one as the first five principal component factors extracted from individual stock returns and the other as the Fama-French three factors. They find that both the size and book-to-market effects are very strong in the model with principal component factors but become much attenuated in the model with Fama-French three factors. Their main conclusion is that there are many other firm-specific variables, such as trading volume, dividend yield, stock price level, and past returns, not considered by Fama and French (1992) that can explain returns and the Fama-French three factors cannot not explain these additional firm-specific variables very well. Zhang (2009) also considers two sets of factors: the first set is extracted from individual stock returns and the second set is from size- and book-to-market sorted 100 portfolio returns. He arrives at a similar conclusion as Brennan, Chordia, and Subrahmanyam (1998) for the first set of factors but finds that the firm-specific variables are insignificantly different from zero when they use more than 3 principal components for the second set of factors. The size and book-to-market effects can be explained by the betas of the three principal component factors extracted from the one-hundred size- and book-to-market-sorted portfolios, indicating somewhat inconsistency between his result and ours. Such inconsistency may arise for two reasons: one is that Zhang (2009) does not allow unobserved factors in his model and only considers the use of various constructed factors, and the other is that Zhang's (2009) sampling period is different from ours. With time-varying betas for factors and firm-specific variables, Chordia, Goyal, Shanken (2012) find that firm characteristics such as size, book-to-market ratio, and the lagged 6-month returns are significant for both the CAPM and Fama-French three-factor models. Our results are consistent with their conclusion even though we use a different framework.

5.3.2 Test for the linearity of $g(\cdot)$

Table 3 reports the bootstrap p-values for the tests of linearity of $g(\cdot)$ in various factor pricing models. For the full sample, all the p-values for all the tests are smaller than 10% for all factor pricing models except for the pure unobserved factor model (PUF) where the information criteria tend to choose a larger number of unobserved factors than in other models. We can reject the null of linearity for the models that include either one observed factor (CAPM) or more than one observed factors (FF, 4-factor and 5-factor models).

For the first subsample, when the test for anomaly effects fails to detect anomalies in FF, 4-factor, and 5-factor models, our test of linearity also does so even at the 10% significance level as expected. However, for the pure unobserved factor model and CAPM, there are no uniform results: our test rejects the null of linearity for some choices of numbers of knots ($q = 3, 4, 5$ in PUF, and $q = 5, 6, 8$ in CAPM) at the 10% significance level and fails to do so for other choices of number of knots. For the second and third subsamples, the test displays a similar pattern as the full sample case. The p-values for all the tests are larger than 10% in PUF and are all smaller than 5% in FF, 4-factor, and 5-factor models. We cannot reject the null of linearity in PUF at the 10% significance level and can readily reject the null of linear at the 5% significance level for in FF, 4-factor, and 5-factor models. Again, the linearity test does not deliver uniform results in CAPM since it rejects the null of linearity at the 10% significance level only when $q = 3, 4, 5$.

In short, when the test for anomaly effects does not provide evidence of anomaly effects, our test of linearity also provides support on the linear null hypothesis; when the test for anomaly effects reveals the presence of anomaly effects, our test of linearity indicates that in most cases the linearity of $g(\cdot)$ can also be rejected at least for some choices of q in CAPM and for all choices of q in FF, 4-factor, and 5-factor models at the 5% level. This implies that the linear functional form tends to be incorrectly specified when anomaly effects are present in these commonly used factor pricing models.

Table 3: Bootstrap p-values for the test of linearity of g(.) in various asset pricing models

	$H_0^{(l)}: \Pr\left[g\left(X_{it}\right)=X_{it}'\theta^0\right]=1$ for some $\theta^0\in\Theta$						
	R	$J_{s,3}$	$J_{s,4}$	$J_{s,5}$	$J_{s,6}$	$J_{s,7}$	$J_{s,8}$
1973/07-2009/12 (full sample)							
PUF	5	0.785	0.780	0.968	0.963	0.993	0.999
CAPM	4	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
FF	3	0.002	<0.001	<0.001	<0.001	<0.001	<0.001
4F	3	0.005	0.002	0.001	0.001	0.002	0.002
5F	3	0.055	0.018	0.043	0.049	0.059	0.037
1973/07-1983/12							
PUF	4	0.073	0.081	0.069	0.193	0.241	0.256
CAPM	4	0.113	0.106	0.075	0.079	0.110	0.063
FF	2	0.243	0.253	0.244	0.246	0.209	0.248
4F	2	0.654	0.701	0.616	0.543	0.387	0.425
5F	2	0.446	0.532	0.460	0.438	0.319	0.317
1984/01-1996/12							
PUF	4	0.250	0.309	0.766	0.813	0.871	0.537
CAPM	3	0.060	0.073	0.096	0.128	0.160	0.173
FF	1	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
4F	1	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
5F	1	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
1997/01-2009/12							
PUF	7	0.952	0.421	0.614	0.739	0.799	0.653
CAPM	7	0.014	0.007	0.088	0.120	0.201	0.245
FF	4	0.001	0.001	0.003	0.003	0.002	0.002
4F	4	0.004	0.001	0.007	0.010	0.004	0.005
5F	4	0.004	0.002	0.008	0.010	0.002	0.005

Note: $J_{s,q}$ (q=3, 4,...,8) denote the linearity test based on the L_2-distance between the sieve and linear estimates.

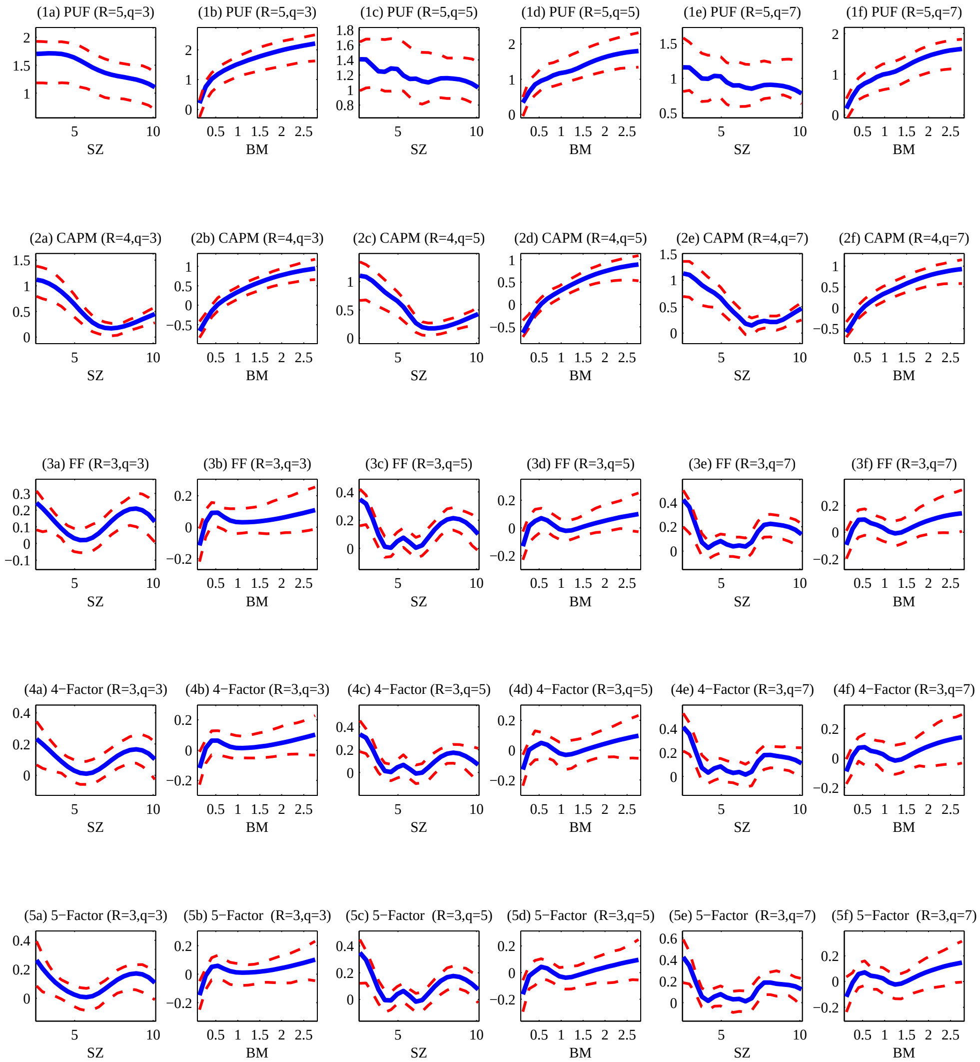

Figure 1: Relationship between excess returns and size (SZ) and book-to-market ratio (BM) in the full sample (1973/07-2009/12). First row: Pure unobserved factor model (PUF), second row: CAPM, third row: Fama-French three-factor model (FF), fourth row: FF+Mom model (4-Factor), and fifth row: FF+Mom+Liq model (5-Factor). Y-axis: excess returns, solid lines: estimated function g(·), dash lines: bootstrap 90% confidence bands.

5.4 Estimation results

To illustrate our estimation results, for the full sample, we plot the estimate $\hat{g}(\cdot,\cdot)$ against each of its two arguments when the other is fixed at its sample mean for the full sample in Figure 1. Denote $\overline{SZ}$ and $\overline{BM}$ as sample means for the size and book-to-market ratio, respectively. The first and second columns respectively report the estimates of $g\left(SZ,\overline{BM}\right)$ and $g(\overline{SZ},BM)$ using 3 knots in the sieve approximation, the third and fourth columns report the above estimates of using 5 knots in the sieve approximation, and the last two columns report the above estimates using 7 knots in sieve approximation. The five rows correspond to the estimates of $g\left(SZ,\overline{BM}\right)$ and $g(\overline{SZ},BM)$ based on the five factor pricing models discussed above. For example, the first row plots the estimated functions in PUF, and the second row plots the estimated functions in CAPM, where the 90% pointwise bootstrap confidence bands are also provided.

We summarize some important findings from Figure 1. First, we can see that all the estimated functions are significantly different from zero at the 10% level, which confirms the test results in Table 2. Second, we can see a significant nonlinear pattern in CAPM, FF, 4-factor and 5-factor models and a rough linear pattern in PUF, which confirms the test results in Table 3. For the first, third and fifth columns, the estimated functions appear to decrease linearly in PUF and alter nonlinearly in CAPM, FF, 4-factor and 5-factor models as the size increases. From the second, fourth, and sixth columns, the estimated functions appear to increase roughly in PUF and CAPM but with a decreasing rate as the book-to-market ratio increases; however, as the book-to-market ratio increases, the estimated functions tend to increase first then start to decrease at around 0.47 and finally increase at a decreasing rate. Third, the estimation results are kind of robust to different choices of numbers of sieve terms. As more terms are used in the sieve approximation, the estimated functions become more zigzag as expected but they still have similar shapes.

Table 4 reports the sample mean and standard deviation of the estimated factor loadings. The table gives results on linear estimation and sieve estimation with 3, 5, and 7 knots used in the construction of sieve bases. In PUF, we only report the results for the five unobserved factor loadings; in CAPM, we report the results for MKT and the four unobserved factor loadings; in FF, 4-factor and 5-factor models, the results are available for the factor loadings of both the observed/constructed 3, 4 or 5 factors and the three unobserved factors. It is worth mentioning that the reported values indicate the sample mean and standard deviation of the *estimated* factor loadings. As $T\to\infty$, elements of the estimated factor loadings converge to their true values at the usual $\sqrt{T}$-rate and follow the asymptotic normal distribution so that we can rely on the sample mean and standard deviations to test whether the population mean is zero or not. A value of the sample mean over the standard deviation that is larger than 1.96 (2.576) indicates the population mean is significantly different from zero at the 5% (1%) level. Interestingly, only the population mean of the factor loadings of Factor 1 in PUF and that

Table 4: Estimation results for factor loadings in the full sample 1973/07-2009/12: the sample average and standard deviation

	R	MKT	SMB	HML	Mom	Liq	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
PUF											
linear	5	-	-	-	-	-	5.163	0.243	0.035	0.015	0.026
							(0.981)	(1.589)	(1.144)	(0.709)	(0.601)
q=3	5	-	-	-	-	-	5.191	0.235	0.045	0.014	0.027
							(0.972)	(1.596)	(1.134)	(0.708)	(0.601)
q=5	5	-	-	-	-	-	5.191	0.237	0.041	0.014	0.027
							(0.972)	(1.596)	(1.121)	(0.708)	(0.601)
q=7	5	-	-	-	-	-	5.178	0.240	0.038	0.014	0.026
							(0.977)	(1.593)	(1.116)	(0.709)	(0.603)
CAPM											
linear	4	1.064	-	-	-	-	1.855	0.471	0.190	0.075	-
		(0.161)					(1.315)	(1.488)	(0.741)	(0.679)	
q=3	4	1.057	-	-	-	-	1.838	0.491	0.203	0.070	-
		(0.164)					(1.339)	(1.462)	(0.777)	(0.676)	
q=5	4	1.058	-	-	-	-	1.837	0.491	0.202	0.070	-
		(0.164)					(1.338)	(1.462)	(0.774)	(0.676)	
q=7	4	1.058	-	-	-	-	1.838	0.492	0.202	0.070	-
		(0.163)					(1.339)	(1.462)	(0.774)	(0.676)	
FF											
linear	3	1.022	0.550	0.300	-	-	0.513	0.235	0.304	-	-
		(0.095)	(0.461)	(0.417)			(0.758)	(0.663)	(0.558)		
q=3	3	1.021	0.550	0.297	-	-	0.518	0.254	0.292	-	-
		(0.096)	(0.460)	(0.418)			(0.755)	(0.659)	(0.569)		
q=5	3	1.022	0.550	0.300	-	-	0.518	0.262	0.288	-	-
		(0.095)	(0.461)	(0.417)			(0.754)	(0.657)	(0.573)		
q=7	3	1.021	0.550	0.297	-	-	0.517	0.267	0.286	-	-
		(0.096)	(0.460)	(0.418)			(0.755)	(0.655)	(0.575)		
4-Factor											
linear	3	1.014	0.551	0.290	-0.036	-	0.336	0.185	0.481	-	-
		(0.087)	(0.461)	(0.415)	(0.061)		(0.535)	(0.651)	(0.771)		
q=3	3	1.013	0.550	0.288	-0.037	-	0.332	0.200	0.482	-	-
		(0.089)	(0.460)	(0.416)	(0.061)		(0.542)	(0.647)	(0.770)		
q=5	3	1.013	0.551	0.289	-0.037	-	0.330	0.208	0.482	-	-
		(0.089)	(0.461)	(0.417)	(0.061)		(0.556)	(0.646)	(0.770)		
q=7	3	1.013	0.551	0.289	-0.038	-	0.330	0.212	0.481	-	-
		(0.089)	(0.461)	(0.417)	(0.061)		(0.547)	(0.645)	(0.771)		
5-Factor											
linear	3	1.013	0.552	0.292	-0.035	0.005	0.334	0.189	0.480	-	-
		(0.088)	(0.461)	(0.416)	(0.061)	(0.158)	(0.534)	(0.649)	(0.771)		
q=3	3	1.011	0.550	0.290	-0.035	0.007	0.323	0.203	0.479	-	-
		(0.090)	(0.460)	(0.417)	(0.061)	(0.193)	(0.549)	(0.646)	(0.772)		
q=5	3	1.011	0.552	0.290	-0.036	0.007	0.328	0.209	0.478	-	-
		(0.090)	(0.461)	(0.417)	(0.061)	(0.193)	(0.541)	(0.645)	(0.771)		
q=7	3	1.011	0.552	0.290	-0.037	0.007	0.328	0.212	0.478	-	-
		(0.090)	(0.460)	(0.417)	(0.061)	(0.193)	(0.543)	(0.644)	(0.772)		

Note: (i) Factors 1-5 represent the estimated unobserved factors; (ii) The values in the parentheses are the sample standard deviation of estimated factor loadings; (iii) q denotes the number of knots in the sieve estimation.

Table 5: Estimation results for factor risk premium in the full sample 1973/07-2009/12

	R	MKT	SMB	HML	Mom	Liq
CAPM						
linear	4	0.3396 (0.2209)	–	–	–	–
$q=3$	4	0.0478 (0.2226)	–	–	–	–
$q=5$	4	0.0559 (0.2224)	–	–	–	–
$q=7$	4	0.0543 (0.2224)	–	–	–	–
FF						
linear	3	0.3916 (0.2205)	0.2009 (0.1579)	0.5111 (0.1559)	–	–
$q=3$	3	0.3592 (0.2205)	0.1985 (0.1579)	0.5099 (0.1559)	–	–
$q=5$	3	0.3570 (0.2205)	0.1982 (0.1579)	0.5099 (0.1559)	–	–
$q=7$	3	0.3555 (0.2205)	0.1979 (0.1579)	0.5099 (0.1558)	–	–
4-Factor						
linear	3	0.4119 (0.2205)	0.2079 (0.1579)	0.5174 (0.1558)	1.1565 (0.4024)	–
$q=3$	3	0.3837 (0.2205)	0.2056 (0.1579)	0.5160 (0.1558)	1.1427 (0.4026)	–
$q=5$	3	0.3792 (0.2205)	0.2059 (0.1579)	0.5159 (0.1558)	1.1407 (0.4026)	–
$q=7$	3	0.3791 (0.2205)	0.2056 (0.1579)	0.5159 (0.1558)	1.1435 (0.4025)	–
5-Factor						
linear	3	0.4244 (0.2205)	0.2119 (0.1579)	0.5110 (0.1558)	1.1623 (0.4024)	1.4769 (0.7584)
$q=3$	3	0.3892 (0.2205)	0.2103 (0.1579)	0.5097 (0.1558)	1.1439 (0.4026)	1.4432 (0.7578)
$q=5$	3	0.3856 (0.2205)	0.2108 (0.1579)	0.5097 (0.1558)	1.1421 (0.4026)	1.4406 (0.7563)
$q=7$	3	0.3841 (0.2205)	0.2111 (0.1579)	0.5095 (0.1558)	1.1437 (0.4024)	1.4488 (0.7560)

Note: The values in the parentheses are the estimated standard deviations of the estimated risk premia.

of MKT in all factor pricing models with observed factors are significantly different from 0 at the 1% level. In addition, we notice that the sample means and standard deviations remain quite stable for the sieve estimates with different choices of numbers of knots (q) and there are only slightly differences between the linear and sieve estimation.

Once we obtain the estimates $\hat{g}(X_{it})$ (where $X_{it} = (SZ_{it}, BM_{it})'$, $i = 1, ..., N$, $t = 1, ..., T$), we can follow Zhang (2009) and estimate the monthly risk premium. The procedure goes as follows. First we define new excess returns after removing the effects of individual asset-specific characteristics: $Y_{it}^* = Y_{it} - \hat{g}(X_{it})$. Given the set of factors $\{F_t^*\}$ that include both the observed/constructed factors and the estimated unobserved factors, we can apply the two-pass method to $\{Y_{it}^*\}$ and $\{F_t^*\}$ to estimate the risk premiums. Table 5 reports the results for the estimated factor risk premium for the observed factors in different factor pricing models. Note that the sieve estimates based on different numbers of knots are quite close to each other, and they are also quite close to the linear estimates for all factors other than MKT in FF, 4-factor, and 5-factor models. In addition, the linear estimates tend to over-estimate the risk premium for all observed common factors and the difference between the linear and sieve estimates of MKT risk premium is huge in CAPM.

6 Conclusion

In this paper we propose a new framework to test for anomaly effects in different asset pricing models. The new setup is quite flexible: it allows for the presence of both observed and unobserved factors and uses a nonparametric function to capture anomaly effects for some asset-specific characteristics. We find strong evidence of the presence of anomaly effects for the size and book-to-market ratio in factor pricing models and when anomaly effects are present, they tend to enter the model nonlinearly.

References

Andrews D (1991) Asymptotic optimality of generalized C_L, cross-validation, and generalized cross-validation in regression with heteroskedastic errors. Journal of Econometrics 47(2): 357–377.

Ang A, Hodrick RJ, Xing Y, Zhang X (2006) The cross-section of volatility and expected returns. Journal of Finance 61(1): 259–299.

Ang A, Kristensen D (2012) Testing conditional factor models. Journal of Financial Econometrics 106(1): 132–156.

Ahn SC, Gadarowski C, Perez MF (2012) Robust two-pass cross-sectional regressions: a minimum distance approach. Journal of Financial Econometrics 10(4): 669–701.

Bai J (2009) Panel data model with interactive fixed effects. Econometrica 77(4): 1229–1279.

Bai J, Ng S (2002) Determining the number of factors in approximate factor models. Econometrica 70(1): 191–221.

Banz RW (1981) The relationship between return and market value of common stocks. Journal of Financial Economics 9(1): 3–18.

Basu S (1977) Investment performance of common stocks in relation to their price-earnings ratios: a test of the efficient market hypothesis. Journal of Finance 32(3): 129–156.

Belloni A, Chernozhukov V (2013) Least squares after model selection in high-dimensional sparse models. Bernoulli 19(2): 521–547.

Berk JB (1995) A critique of size-related anomalies. Review of Financial Studies 8(2): 275–286.

Black F, Jensen MC, Scholes M (1972) The capital asset pricing model: some empirical tests. In Jensen MC (ed), Studies in the Theory of Capital Markets, Praeger, New York.

Brennan M, Chordia T, Subrahmanyam A (1998) Alternative factor specifications, security characteristics, and the cross-section of expected stock returns. Journal of Financial Economics 49(3): 345–373.

Carhart MM (1997) On persistence in mutual fund performance. Journal of Finance 52(1): 57–82.

Chamberlain G, Rothschild M (1983) Arbitrary, factor structure, and mean-variance analysis on large asset markets. Econometrica 51(5): 1281—1304.

Chen N, Roll R, Ross SA (1986) Economic forces and the stock market. Journal of Business 59(3): 383–404.

Chen X (2007) Large sample sieve estimation of semi-nonparametric models. In Heckman JJ, Leamer E (eds), Handbook of Econometrics 6: 5549–5632, North Holland, Amsterdam.

Chen X (2013) Penalized sieve estimation and inference of semi-nonparametric dynamic models: a selective review. In Acemoglu D, Arellano M, Dekel E (eds), Advances in Economics and Econometrics 3: 485–544, Cambridge University Press, Cambridge.

Chordia T, Goyal A, Shanken J (2012) Cross-sectional asset pricing with individual stocks: betas versus characteristics. Working Paper, Emory University.

Chordia T, Shivakumar L (2002) Momentum, business cycle and time-varying expected returns. Journal of Finance 57(2): 985–1019.

Cochrane J (2005) Financial markets and the real economy. NBER Working Paper 11193.

Connor G, Hagmann M, Linton O (2012) Efficient semiparametric estimation of the Fama–French model and extensions. Econometrica 80(2): 713–754.

Connor G, Korajczyk RA (1986) Performance measurement with the arbitrage pricing theory: a new framework for analysis. Journal of Financial Economics 15(3): 373–394.

Connor G, Korajczyk RA (1988) Risk and return in an equilibrium APT: application of a new testing methodology. Journal of Financial Economics 21(2): 255-289.

Connor G, Korajczyk RA (1993) A test for the number of factors in an approximate factor model. Journal of Finance 48(4): 1263–1291.

Connor G, Linton O (2007) Semiparametric estimation of a characteristic-based factor model of common stock returns. Journal of Empirical Finance 14(5): 694–717.

Fama EF, French KR (1992) The cross-section of expected stock returns. Journal of Finance 47(2): 427–465.

Fama EF, French KR (1993) Common risk factors in the returns on stocks and bonds. Journal of Financial Economics 33(3): 3–56.

Fama EF, French KR (1996) Multifactor explanations of asset-pricing anomalies. Journal of Finance 51(1): 55–84.

Fama EF, MacBeth JD (1973) Risk return and equilibrium: empirical tests. Journal of Political Economy 81(3): 607–636.

Friend I, Blume M (1970) Measurement of portfolio performance under uncertainty. American Economic Review 60(4): 561–575.

Goyal A (2012) Empirical cross-sectional asset pricing: a survey. Financial Markets and Portfolio Management 26(1): 3–38.

Grauer RR (1999) On the cross-sectional relation between expected return, betas and size. Journal of Finance 54(2): 773–789.

Härdle W, Mammen E (1993) Comparing nonparametric versus parametric regression fits. Annals of Statistics 21(4): 1926–1947.

Hong Y, White H (1995) Consistent specification testing via nonparametric series regression. Econometrica 63(5): 1133–1159.

Huberman G (1982). A simple approach to arbitrage pricing theory. Journal of Economic Theory 28(1): 183-191.

Jagannathan R, Skoulakis G, Wang Z (2003) Testing linear asset pricing factor models using cross-sectional regressions and security characteristics: asymptotic theory. Working Paper, Northwestern University.

Jagannathan R, Skoulakis G, Wang Z (2010) The analysis of the cross-section of security returns. In Aït-Shahalia Y, Hansen LP (eds.) Handbook of Financial Econometrics, Vol 2 Applications: 73-134.

Jagannathan R, Wang Z (1998) An asymptotic theory for estimating beta-pricing models using cross-sectional regression. Journal of Finance 53(4): 1285–1309.

Jegadeesh N, Titman S (1993) Returns to buying winners and selling losers: implications for stock market efficiency. Journal of Finance 48(1): 65–91.

Kleibergen K, Zhan Z (2013) Unexplained factors and their effects on second pass R-squared's and t-tests. Working Paper, Brown University.

Lehmann BN, Modest DM (1988) The empirical foundations of the arbitrage pricing theory. Journal Financial Economics 21(2): 213–254.

Lewellen J, Nagel S, Shanken J (2010) A skeptical appraisal of asset-pricing tests. Journal of Financial Economics 96(2): 175–194.

Li Y, Yang L (2011) Testing conditional factor models: a nonparametric approach. Journal of Empirical Finance 18(5): 972–992

Li Y, Su L, Xu Y (2013) A combined approach to the inference of conditional factor models. Working Paper, Department of Finance, Temple University.

Lin H, Hong Y (2006) New test of asset pricing models in China. Working Paper, WISE, Xiamen University.

Lintner J (1965) The valuation of risk assets and selection of risky investments in stock portfolios and capital budgets. Review of Economics and Statistics 47(1): 13–37.

Lu X, Su, L. (2014) Shrinkage estimation of dynamic panel data models with interactive fixed effects. Working Paper, School of Economics, Singapore Management University.

Markowitz H (1952) Portfolios selection. Journal of Finance 7(1): 77–91.

Moon HR, Weidner M (2013a) Dynamic linear panel data regression models with interactive fixed effects. Working Paper, University of Southern California.

Moon HR, Weidner M (2013b) Linear regression for panel with unknown number of factors as interactive fixed effects. Working Paper, University of Southern California.

Pástor L, Stambaugh, RF (2003) Liquidity risk and expected stock returns. Journal of Political Economy 111(3): 642–685.

Rosenberg B, Reid K, Lanstein R (1985) Persuasive evidence of market inefficiency. Journal of Portfolio Management 11(3): 9–16.

Ross, SA (1976) The arbitrage theory of capital asset pricing. Journal of Economic Theory 13(3): 341–360.

Subrahmanyam A (2010) The cross-section of expected stock returns: what have we learnt from the past twenty-five years of research? European Financial Management 16(1): 27–42.

Serlenga L, Shin Y, Snell A (2004) A panel data approach to testing anomaly effects in factor pricing models. Discussion Paper 88, Edinburgh School of Economics, University of Edinburgh.

Shanken J (1992) On the estimation of beta-pricing models. Review of Financial Studies 5(1): 1–33.

Sharpe W (1964) Capital asset prices: a theory of market equilibrium under conditions of risk. Journal of Finance 19(3): 425–442.

Stambaugh RF (1982) On the exclusion of assets from tests of the two-parameter model: a sensitivity analysis. Journal of Financial Economics 10(3): 237–268.

Su L, Chen Q (2013) Testing homogeneity in panel data models with interactive fixed effects. Econometric Theory 29(6): 1079–1135.

Su L, Jin S, Zhang Y (2013) Specification test for panel data models with interactive fixed effects. Working Paper, Singapore Management University.

Su L, Lu X (2013) Nonparametric dynamic panel data models: kernel estimation and specification testing. Journal of Econometrics 176(2): 112–133.

Su L, Zhang Y (2013) Nonparametric dynamic panel data models with interactive fixed effects: sieve estimation and specification testing. Working Paper, Singapore Management University.

Zhang, C (2009) On the explanatory power of firm-specific variables in cross-sections of expected returns. Journal of Empirical Finance 16(2): 306–317.